NBER WORKING PAPER SERIES

AIR POLLUTION AND INTERNAL MIGRATION IN THE UNITED STATES

Michael Keller
Christopher R. Knittel
Benjamin Krebs
Simon Luechinger

Working Paper 35317
http://www.nber.org/papers/w35317

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2026

For comments and suggestions, we thank all participants at the Allied Social Science Associations Annual Meeting in San Antonio, TX (January 2024), the Swiss Society of Economics and Statistics Annual Congress (June 2024), the AURÖ Workshop for Young Researchers in Environmental and Resource Economics (February 2025), and the Zurich Environmental Economics Workshop (March 2026), as well as at the research seminars of the University of Lucerne (November 2024) and the University of Maastricht (April 2025). Benjamin Krebs acknowledges funding by the Swiss National Science Foundation Grant Number PZ00P1_223786. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Air Pollution and Internal Migration in the United States
Michael Keller, Christopher R. Knittel, Benjamin Krebs, and Simon Luechinger
NBER Working Paper No. 35317
June 2026
JEL No. Q53, R23

ABSTRACT

We estimate the effect of $PM_{2.5}$ pollution on migration between commuting zones in the United States from 2005-2019. To account for the correlation between origin and destination commuting zones' pollution levels and potential endogeneity, we estimate a dyadic migration model and isolate permanent changes in origin and destination pollution emanating from distant coal-fired power plants. Annual panel and long-difference estimates indicate that air pollution plays a key role in relocation decisions. For the typical commuting zone, an isolated average 2005-2019 $PM_{2.5}$ concentration decrease of 3.85 μg/m^3 would avert out-migration and increase in-migration, totaling 2 percent of the population annually.

Michael Keller
University of Lucerne
michael.keller@unilu.ch

Christopher R. Knittel
Massachusetts Institute of Technology
MIT Sloan School of Management
Department of Economics
and NBER
knittel@mit.edu

Benjamin Krebs
University of Basel
Faculty of Business and Economics
benjamin.krebs@unibas.ch

Simon Luechinger
University of Lucerne
Simon.Luechinger@unilu.ch

1 Introduction

Large swathes of the U.S. population worry about exposure to local air pollution: According to a 2024 Gallup survey, 40 percent of respondents express moderate or strong concerns.[1] Although relocation is a consequential decision, moving to a cleaner area constitutes the only viable solution to lower exposure to ambient air pollution in the long term. Indeed, the same survey suggests that people choose this solution: A full three percent of respondents claim to have permanently moved due to environmental pollution in the twelve months preceding the survey. However, there is a dearth of systematic evidence on how air pollution affects internal migration in the United States. Such evidence is crucial for understanding how individuals adapt to poor environmental conditions in the long term and the importance they place on clean air.

We ask if U.S. households move in response to air pollution. We study migration flows across commuting zones in the United States and the years 2005-2019, leveraging IRS data together with annual panel and long-difference instrumental variables estimators. This period saw substantial, persistent air quality improvements across large parts of the country. Since migration decisions arguably depend on long-term expectations rather than transient fluctuations, this is an ideal setting to analyze the internal migration effects of air pollution.

Estimating causal relocation responses to air pollution presents two challenges. First, pollution-generating activities may affect migration directly, and local pollution levels are often poorly measured (e.g., Chay and Greenstone 2003 on the endogeneity of pollution). To estimate causal migration responses, we isolate long-term fluctuations in air pollution from changes in emissions at distant coal-fired power plants in an instrumental variable setting. During our sample period, power plant emissions fell substantially because electricity generation from natural gas and renewables largely displaced coal and because of the continuing adoption of emission controls (Knittel et al. 2015; Fell and Kaffine 2018; Johnsen et al. 2019; Coglianese et al. 2020). Second, since people disproportionately move over shorter distances, changes in pollution between migration-linked commuting zones are correlated. Ignoring these correlated pollution changes biases estimates toward zero (Borusyak et al. 2022). Therefore, we estimate models for directed dyads and account for pollution in both the origin and destination commuting zones.

We find that the shares of moving households and individuals increase with fine particulate matter ($PM_{2.5}$) pollution in the origin commuting zone and decrease with $PM_{2.5}$ pollution in the destination commuting zone. This general pattern only emerges for between-state relocations, not for within-state relocations. The relative importance of origin and destination pollution slightly differs between estimation approaches. Using the annual panel data, the relocation response to destination pollution is smaller than the one for origin pollution, and it is statistically insignificant. The "push effects" of pollution suggest that residents are aware of pollution levels in their commuting zone. In complementary analyses, we document positive associations between self-reported worries about air pollution and actual pollution levels. With the long-difference estimates for the

[1] https://news.gallup.com/poll/643604/black-adults-concerned-others-local-pollution.aspx, accessed December 19, 2025.

years 2005 and 2019, we again detect positive effects for origin pollution and negative effects for destination pollution. In contrast to the annual panel estimates, the effects for destination pollution are (slightly) larger than the effects for origin pollution, and they are statistically significant. In a quantification exercise based on the long-difference estimates, we illustrate the importance of pollution for relocation decisions: An isolated reduction in $PM_{2.5}$ pollution of 3.85 µg/m^3, corresponding to the average observed decrease over our sample period, would lead to annual avoided out-migration and additional in-migration totaling about 2 percent of the commuting zone population. However, since changes in air quality are typically not isolated but instead affect origin and destination commuting zones similarly, observed pollution-induced migration will be lower.

Four different literatures are relevant to our analysis. First and most directly, we contribute to a small literature on air pollution and migration. Chen et al. (2022) and Khanna et al. (2025) document strong in- and out-migration responses to inversion-induced transient fluctuations in air pollution in China. Khanna et al. (2025) further report similar migration effects in response to more permanent shifts in pollution resulting from changes in coal consumption at distant power plants. Similarly, according to Mikula and Pytliková (2026), highly-polluted municipalities in North Bohemia (Czech Republic) experienced an increase in the net migration rate following the desulfurization of power plants. There is little evidence on the pollution-migration relationship for the United States. Borgschulte et al. (2024) and De Silva et al. (2024), whose main focus lies on questions related to pollution and income, provide evidence on migration in their respective appendices. Borgschulte et al. (2024) find no effects of wildfire-induced transient air pollution changes on either in- or out-migration. In contrast, for more permanent shifts caused by environmental policies, De Silva et al. (2024) document negative effects of air pollution on counties' net migration.[2] Using a dyadic migration model with differences between origin and destination pollution, but not accounting for the endogeneity of pollution, they find qualitatively similar results. No prior study simultaneously addresses both the endogeneity of pollution and the migration-linked correlations in pollution. As discussed above, we address both challenges in our analysis for the United States.

Second, according to another literature, air pollution affects location decisions (e.g., Bayer et al. 2009), house prices (e.g., Chay and Greenstone 2005; Fraenkel et al. 2024), and neighborhood composition (Heblich et al. 2021; Currie et al. 2023). This literature implies that people consider air quality in their relocation decisions, but it does not directly quantify the corresponding migration responses.

Third, studies of internal migration emphasize economic determinants such as wages (Sprung-Keyser et al. 2022; Olney and Thompson 2026), housing prices (Olney and Thompson 2026), and fracking-induced opportunities (Bartik et al. 2019; Wilson 2022), but also environmental amenities other than pollution such as local climate conditions (Baylis et al. 2025).

Fourth, fracking-induced declines in natural gas prices lowered coal's importance in electricity generation (Knittel et al. 2015) and led to coal plant retirements (Davis et al. 2022) with concomi-

[2] Related, Banzhaf and Walsh (2008) and De Silva et al. (2024) find that proximity to Toxics Release Inventory sites and emissions reduces local population figures and net migration, respectively.

tant reductions in air pollution (Johnsen et al. 2019; Burney 2020). We exploit these fuel shifts for identification and provide further evidence of their beneficial effects on local air quality.

2 Empirical Strategy

To assess the relocation responses to air pollution, we estimate a dyadic migration model. Specifically, we regress the share of households or individuals who move from an origin commuting zone o to a destination commuting zone d in year t, M_{odt}, on the instrumented $PM_{2.5}$ pollution in the origin and destination commuting zones in year t, $\hat{P}_{ot}$ and $\hat{P}_{dt}$. The unit of analysis is thus the directed dyad per year. We do not include stayers (with $o = d$).

Since relocations are more frequent over short distances, pairs of close commuting zones typically observe larger migration flows than pairs of commuting zones that are far apart. Due to their proximity, these migration-linked commuting zones will also experience similar pollution. Ignoring this correlation would bias the estimated effect of pollution towards zero (Borusyak et al. 2022). Our dyadic model includes both origin and destination pollution and, thus, accounts for this correlation.

We control for directed dyad effects, δ_{od}, to absorb time-invariant factors such as distance or historical migration links. We further add origin state × year and destination state × year effects, $\theta_{s(o)t}$ and $\theta_{s(d)t}$.[3] These state-by-year effects capture broader environmental and economic developments such as environmental policies, fracking booms, or mining slumps. If these state-by-year effects correlate with predicted pollution, our instrumental variable approach (see below) is only valid conditional on their inclusion. Indeed, ignoring these state-by-year effects strongly affects our results and changes the sign of the second-stage estimates. Further, we control for heating- and cooling-degree days and total precipitation in the origin and destination commuting zones in a given year, $\boldsymbol{W}_{ot}$ and $\boldsymbol{W}_{dt}$. Equation 1 summarizes our model for the second stage:

$$M_{odt} = \beta_o \hat{P}_{ot} + \beta_d \hat{P}_{dt} + \boldsymbol{W}'_{ot}\gamma_o + \boldsymbol{W}'_{dt}\gamma_d + \delta_{od} + \theta_{s(o)t} + \theta_{s(d)t} + \varepsilon_{odt} \tag{1}$$

We report two standard errors. The first standard error accounts for two-way clustering at the levels of origin and destination commuting zones (as suggested by Cameron et al. 2011 for dyadic data). The second standard error allows for spatial correlations by moving clustering to the levels of origin and destination U.S. Bureau of Economic Analysis economic areas.

To identify causal relationships, we must account for the fact that economic and other pollution-generating activities have direct effects on the relocation decisions of households and individuals. We address this well-known endogeneity problem by instrumenting pollution in the origin and destination commuting zones with predicted $PM_{2.5}$ pollution due to emissions at distant coal-fired power plants. Using an atmospheric dispersion model, we simulate a commuting zone's $PM_{2.5}$ pollution that emanates from sulfur dioxide (SO_2), nitrogen oxides (NO_x), and $PM_{2.5}$ emissions

[3] We assign commuting zones to the state in which the highest share of the 2020 commuting zone population lived. Three states (Connecticut, Delaware, and Rhode Island) are matched with only one commuting zone. These commuting zones and states have no observations for within-state relocations.

at distant coal plants. We consider plants with a distance from the commuting zone boundary of more than 100 kilometers for our baseline results and vary this distance threshold in robustness analyses.[4]

Let us walk through the construction of this instrument step by step. First, we estimate emission-concentration coefficients with the atmospheric dispersion model on a 25 $\times$ 25 kilometers grid. The emission-concentration coefficient, c_{jpg}, measures the increase in the annual mean $PM_{2.5}$ concentration in a grid cell g due to a one-ton increase in emissions of pollutant j at power plant p. Using the dispersion model InMAP (Tessum et al. 2017), we simulate the grid-level $PM_{2.5}$ concentration from an emitted ton of a specific pollutant, with separate simulations for each of 538 power plants and each pollutant. InMAP is not only capable of modeling atmospheric dispersion but also the chemical transformation of gaseous pollutants into particulate pollutants. Thus, we can use both primary and secondary $PM_{2.5}$ pollution. The InMAP simulation relies on 2005 weather data and requires information on the power plant location, stack height and diameter as well as exhaust velocity and temperature.

Second, we calculate the simulated $PM_{2.5}$ concentration in grid cell g and year t from emissions of pollutant j at power plant p, P^{sim}_{jpgt}, by multiplying emissions, e_{jpt}, with the emission-concentration coefficient: $P^{sim}_{jpgt} = e_{jpt}\, c_{jpg}$.[5] Third, we sum across pollutants and obtain the overall simulated $PM_{2.5}$ concentration at the grid level, $P^{sim}_{pgt} = \sum_j P^{sim}_{jpgt}$. Fourth, we aggregate this simulated $PM_{2.5}$ concentration to the commuting zone level using a weighted average of all grid cells falling into the commuting zone boundaries, whereby we use the share of a commuting zone z's area covered by a grid cell g as weights, ω_{zg}. The simulated $PM_{2.5}$ concentration in commuting zone z and year t from emissions at power plant p is $P^{sim}_{pzt} = \sum_g \omega_{zg}\, P^{sim}_{pgt}$.

Finally, we add the simulated $PM_{2.5}$ concentrations across all power plants with a distance to the commuting zone of more than 100 kilometers. Thus, the simulated $PM_{2.5}$ concentration in commuting zone z and year t is: $P^{sim}_{zt} = \sum_p 1\,\{\text{distance}_{pz} > 100 \text{ kilometers}\}\, P^{sim}_{pzt}$. Figure A1 in the Online Appendix illustrates the distance threshold. It depicts the simulated $PM_{2.5}$ concentration in 2005 originating from one specific power plant and a 100-kilometer-radius circle centered on the plant. In our aggregation, we ignore the simulated concentration for all commuting zones intersecting the circle area, i.e., the ones with thick border lines in the figure. The figure reveals that included commuting zones bordering the excluded ones are usually substantially more than 100 kilometers away from the power plant.

Figure 1 presents the actual (Panels A and B) and simulated (Panels C and D) $PM_{2.5}$ concentrations in the years 2005 (Panels A and C) and 2019 (Panels B and D), together with the coal-fired power plants operating in the respective years. For illustration purposes, we aggregate simulated $PM_{2.5}$ concentrations across all power plants and do not exclude the concentrations originating from nearby plants. As the figure demonstrates, the pollution emanating from coal plants accounts for a substantial portion of the observed spatial and temporal variation in pollution. A third of the

[4]The identification strategy resembles that of Bayer et al. (2009), who model pollution from sources with a distance of more than 80 kilometers.

[5]While InMap models the dispersion of sulfur oxides, SO_x, emissions are reported in SO_2.

decrease in simulated $PM_{2.5}$ pollution is attributable to coal plant retirements: On average, commuting zones' simulated $PM_{2.5}$ concentration (without a distance threshold) falls by 3.30 µg/m^3 between 2005 and 2019. The average reduction attributable to plant closures is 1.09 µg/m^3.[6] The residual changes are, among other factors, attributable to variation in utilization rates, fuel choice, and control technologies, as well as to the opening of new plants.

3 Data

In the following, we present the data sources and detail the construction of the migration, observed and simulated $PM_{2.5}$ concentrations, and control variables used in our analysis.

Migration The migration data are from the Internal Revenue Service (IRS) Statistics of Income Division.[7] The IRS infers migration flows from address changes of tax filers in two consecutive years t and $t+1$. The data capture relocations between two tax seasons, which last from the beginning of the year to mid-April. Thus, most relocations probably occur in year t and, consequently, we assign the migration flows between the years t and $t+1$ to the environmental situation in year t. The number of tax returns and exemptions approximately corresponds to the relocations by households and individuals, respectively. The data ignore non-taxpayers.

The IRS publishes county-level in- and out-migration figures, which we use to construct county-to-county migration flows.[8] Migration flows of less than 10 households (until 2012/2013) or 20 households (from 2013/2014) are omitted, but the corresponding relocations are aggregated into flows to other counties in the same state, other counties in different states, and other counties in the four regions Northeast, Midwest, South, and West. Additional migration flows are sometimes omitted to prevent researchers from recovering omitted figures from aggregated flows. We impute omitted migration flows in the following way. First, we equally allocate the total migration flows to other counties in the same state to all same-state counties with omitted figures. Second, we equally distribute the total migration flows to other counties in one of the four regions to different-state counties in these regions with missing migration figures. If the migration figures for some of the four regions are missing, we first equally distribute the total flows to other states, less the total of the observed flows to other regions, to the regions with missing migration figures. There are 23 counties (84 observations) with missing own flows (with $o = d$). To calculate household and individual totals in an origin location, we impute these missing values with the county's average own flows. We aggregate the county-to-county flows to commuting zone-to-commuting zone flows within the contiguous United States based on Fowler (2024).

[6] To estimate reductions from plant retirements, we aggregate the simulated 2005 pollution across all power plants that were operational in 2005, but not in 2019.

[7] https://www.irs.gov/statistics/soi-tax-stats-migration-data; accessed October 31 and November 2, 2022.

[8] In- and out-migration flows allow us to construct two separate county-to-county migration flows. We construct both variants to reduce the number of missing values. In the very few cases with discrepancies in migration flows across the two variants, we average the respective migration flows.

Our dependent variables are relocating households or individuals in a directed dyad and year expressed as the share of the origin commuting zone's number of households or individuals at the beginning of the year. The relevant start-of-year numbers are the sum of households or individuals across all directed dyads of the origin commuting zone, including the one for stayers, in a given year.

Pollution We use data on annual average $PM_{2.5}$ concentrations from Shen et al. (2024). They use satellite-based pollution measurements and atmospheric dispersion modelling to allocate pollution vertically, and then calibrate the resulting estimates against air quality monitors (see also Van Donkelaar et al. 2021). We aggregate the pollution data from 0.01×0.01 degree grid cells to the commuting zone level in the same way as the simulated pollution (see above).

Instrument To create our instrument, we require data on the locations of power plants, their stacks, and their emissions of the $PM_{2.5}$ precursors SO_2 and NO_x. Further, we need information on firing and control technology configurations, coal consumption and ash content, and emission factors to estimate their $PM_{2.5}$ emissions.

We consider the coal-fired power plants in the EPA's Clean Air Markets Program Data (CAMPD)[9]. For these power plants, we complement the data with information from the Form EIA-860[10] and Form EIA-923[11]. The power plant locations come from CAMPD. For the stack information, we rely on Form EIA-860. For each power plant, we use the information corresponding to the stacks with the maximum height per power plant across all years between 2007 and 2019 recorded in Form EIA-860. Earlier information on stacks is not available. For top area as well as exhaust velocity and temperature, we use the mean per power plant and selected stacks. For power plants that are missing in the EIA data or have no information on stacks, we impute stack information with averages across power plants in the same state or across power plants overall. We obtain facility-level annual SO_2 and NO_x emissions from CAMPD.

To estimate $PM_{2.5}$ emissions, we multiply coal consumption in tons by the relevant emission factor from the EPA[12], which varies with the coal ash content and the power plant firing and PM control technologies. The necessary information is scattered across CAMPD and Form EIA-923 and is not always available for all months and power plants. CAMPD contains total heat input as well as firing and PM control technologies in all years.[13] Form EIA-923 includes heat input and coal

[9] https://campd.epa.gov; accessed March 19, 2024.

[10] https://www.eia.gov/electricity/data/eia860/; accessed October 31, 2024.

[11] https://www.eia.gov/electricity/data/eia923/; accessed October 31, 2024.

[12] https://www.epa.gov/air-emissions-factors-and-quantification/ap-42-fifth-edition-volume-i-chapter-1-external-0; accessed March 3, 2023. We calculate emissions for filterable $PM_{2.5}$, but not condensable $PM_{2.5}$, since the latter are generally much lower.

[13] If CAMPD reports multiple PM control technologies for a boiler and month, we only consider the most effective one according to the EPA document. We use the EPA emissions factors for "scrubber" if the control technology is "wet scrubber" in CAMPD, "multiple cyclones" if it is "cyclone", and "electrostatic precipitator" if it is either "electrostatic precipitator" or "wet ESP". We assume that the control technology "other" has the same effectiveness as the least effective control technology according to the EPA document. If the EPA document, but not the CAMPD, includes emission factors for slightly different technology variants (e.g., cyclones with and without flyash reinjection),

use by coal type in all years and ash content by coal type from 2008.[14] We rely on Form EIA-923 to calculate average heat input per unit of coal and average ash content at the power-plant-by-month level.[15] We combine this information with heat input as well as firing and control technology at the boiler-by-month level from CAMPD.[16] For CAMPD power plants and years not included in the Form EIA-923 or with missing information on heat input per unit and on ash content, we impute the relevant information with the corresponding state-by-year or year average. We divide the CAMPD heat input figures by the EIA heat input per unit of coal to estimate coal consumption in tons.

Primary $PM_{2.5}$ pollution is less important than secondary $PM_{2.5}$ pollution and accounts for only 1.73 percent of simulated $PM_{2.5}$ pollution on average in our instrument.

Weather For weather, we use Schlenker's county-level data[17] based on the fine-grained model data of the PRISM group[18].

4 Results

We find that air pollution in origin commuting zones increases and in destination commuting zones decreases internal migration. The impacts are driven by between-state relocations. With the annual panel data, the effects are stronger for origin than destination pollution. The estimates are unaffected by the choice of the threshold for the distance between commuting zones and power plants, and by the inclusion of house prices and wages. With long-differences, we detect impacts of pollution on internal migration that are of similar magnitudes for origin and destination pollution.

Table 1 reports OLS estimates (Panel A) and instrumental variable estimates (Panel B) from the annual panel data for all relocations (Columns 1 and 2), within-state relocations (Columns 3 and 4), and between-state relocations (Columns 5 and 6), separately for households (Columns 1, 3, and

we use the emission factor for the least effective one. Further, we employ the same emission factors for "dry bottom turbo-fired boiler", "dry bottom vertically-fired boiler", "tangentially-fired", "cell burner boiler", and "arch-fired boiler" as for "dry bottom wall-fired boiler". Similarly, we use the same emission factors of 1.4 for "bubbling fluidized bed boiler" and for "circulating fluidized bed boiler" as for "stoker" with "multiple cyclones with flyash reinjection", not differentiated according to control technology. For boiler types, for which the EPA does not report emission factors, we impute the missing emission factor by state-year-control-technology or year-control-technology averages of power plants with the same PM control technology.

[14]For years before 2008, we impute the ash content per power plant and coal type by the unweighted average ash content across all boilers and months of the year 2008. For power plants and coal types without information on ash content, we instead use the state or overall average ash content in 2008.

[15]For some power plants, heat input per unit of coal and ash content is missing for individual months and coal types. We linearly interpolate these values. We also extrapolate them using the first or last observed value per power plant. For each power plant and month, we compute a weighted average of the heat input per unit of coal and of the ash content across coal types, using the coal-type-specific heat inputs as weights. The EPA reports emission factors for (sub-)bituminous coal and anthracite. We observe no anthracite use in the Form EIA-923. We treat all reported coal types as equivalent to (sub-)bituminous coal.

[16]Even though both CAMPD and Form EIA-923 report the information at the boiler level, we collapse the EIA data to the power-plant level before merging them with CAMPD because we are unable to match boilers across the two datasets.

[17]http://www.columbia.edu/~ws2162/links.html, accessed February 7, 2024.

[18]https://prism.oregonstate.edu/, accessed November 6, 2025

5) and individuals (Columns 2, 4, and 6). The OLS estimates are highly indicative of an omitted variable bias. The coefficients for origin and destination air pollution have a counterintuitive sign. They imply that high air pollution keeps residents in their origin commuting zones but, with the exception of the within-state relocations of individuals, attracts them to destination commuting zones. In the remainder, we focus on the instrumental variable estimates.

Turning to the instrumental variable estimates, we begin our discussion with the two first stages, i.e., one for pollution in the origin commuting zone and one for pollution in the destination commuting zone. In both these first stages, only the corresponding instrumental variable is relevant: simulated $PM_{2.5}$ pollution in the origin commuting zone predicts actual $PM_{2.5}$ concentration in the origin commuting zone and *vice versa* for the destination commuting zone. The relevant coefficients are identical in both first stages and amount to 0.1545 in the models for all relocations. As expected, the opposite instrumental variables have no effects. The tests for overall instrument strength of Lewis and Mertens (2025) vary greatly with the sample and clustering level. They suggest that results for within-state relocations need to be interpreted with caution.

In the reduced form estimates for all relocations and for between-state relocations, the two instrumental variables affect migration flows in the expected direction: origin pollution increases and destination pollution decreases the share of relocating households or individuals. However, only origin pollution has a statistically significant effect. For within-state relocations, the estimates are very imprecise and mostly wrongly signed.

These patterns carry over to the second-stage estimates. The origin-pollution or "push" effects are sizable and statistically significant, the destination-pollution or "pull" effects are smaller (in absolute terms) and statistically insignificant. Thus, in the words of Sprung-Keyser et al. (2022, p. 30; emphasis in original), pollution seems to affect "the choice *whether* to migrate", and less "the choice *where* to migrate". This result may reflect the notions that people's awareness of local pollution and its consequences influence their relocation decisions, and that such awareness more closely aligns with actual pollution at origin locations than at destination locations. However, we should not overemphasize the difference in magnitude of the "push" and "pull" effects, as they are of comparable magnitude in the in the long-difference estimates.

According to the second-stage estimates in Table 1, an increase in $PM_{2.5}$ concentration by 1 $\mu g/m^3$ increases the share of households from the origin commuting zone who move to one particular destination by 0.0044 per 1,000 households and the share of individuals by 0.0058 per 1,000 individuals. While these effects seem small, one should bear in mind that there are 559 possible destinations. We elaborate on the effect magnitudes in more detail when discussing the long-difference estimates. A rise in $PM_{2.5}$ concentration in the destination commuting zones has smaller effects. The magnitude of the "pull effect" is around 73 percent of the magnitude of the "push effect" for households and about 45 percent for individuals.

The patterns just described arise for between-state relocations. Thus, some people who move because of background pollution do so over long distances, possibly because short-distance moves provide limited scope for escaping background pollution.[19] The imprecise estimates for within-state relocations prevent us from drawing any conclusions for short-distance moves.

We find systematic effects of air pollution on households' and individuals' relocation decisions. This suggests that people are aware of environmental conditions and take them into consideration when deciding where to live. Alternatively, they experience pollution's adverse effects (e.g., elevated morbidity or reduced productivity), connect these adverse effects with a location, but not to environmental conditions. Table 2 provides direct evidence for individuals' awareness of their local air pollution. It regresses self-reported subjective concerns about air pollution from the March Gallup Poll Social Series of the years 2006 to 2019 on $PM_{2.5}$ concentrations.[20] Panel A reports the results from OLS, Panel B from instrumental variable estimates. In Columns (1) and (2), we regress the response on a four-point scale on pollution, in Columns (3) and (4), we fit a linear probability model with a dummy for respondents worrying "a great deal".

The OLS estimates in Panel A document a strong association between actual and perceived air pollution. A rise in $PM_{2.5}$ concentration by 1 $\mu g/m^3$ increases the average response by 0.08-0.10 points on the four-point scale and raises the probability of a respondent worrying a great deal by 3.7-4.6 percentage points. In the context of environmental concerns, we are less worried about omitted variables than in the case of migration, though estimates may still be attenuated towards zero because of measurement error. The instrumental variable estimates in Panel B are consistent with this notion. The estimates are about two to five times higher than the OLS estimates, raising the average response by 0.21-0.30 points and the probability of strong concerns by 14.1-18.8 percentage points. However, the instrumental variable estimates are not statistically significant.

We test our results' sensitivity of (i) to changes in the threshold for the distance between commuting zone boundaries and power plants used in constructing the instrument and (ii) to the inclusion of wages and house prices as control variables.

Our baseline results rest on instrumental variables with a 100-kilometer threshold for the inclusion of power plants. We re-estimate with varying thresholds – in steps of 25 kilometers from 0 to 200 kilometers – the first stage for origin actual $PM_{2.5}$ pollution, the reduced form, and the second stage. According to the results in Figure A2 (all relocations), Figure A3 (within-state relocations), and Figure A4 (between-state relocations), we observe very similar patterns to our baseline results independent of the distance threshold. As expected, with increasing distance thresholds, the first-stage coefficient decreases, and the corresponding confidence bands become slightly wider. Further, the prevalence of wrongly signed reduced-form and second-stage estimates for the within-state relocations, which are always statistically insignificant, increases with distance thresholds. Finally, with a 200-kilometer threshold, the estimates for destination pollution are not only statistically insignificant but also close to zero.

Pollution-induced migration affects wages and house prices, which in turn feed back on migra-

[19] We cannot capture very short-distance relocations within commuting zones away from pollution hotspots.

[20] The question reads: "I'm going to read you a list of environmental problems. As I read each one, please tell me if you personally worry about this problem a great deal, a fair amount, only a little, or not at all. First, how much do you personally worry about – [...] B. Air pollution?"

tion. This feedback counteracts the direct impact of pollution on migration. To shut down this feedback and recover the full direct impact, we estimate models with origin and destination wages and house prices as control variables. If the feedback mechanism is important, we would expect the (absolute) value of the pollution coefficients to increase. However, other causal changes are possible. For example, if pollution directly reduces wages through lower productivity (Graff Zivin and Neidell 2012; Chang et al. 2019), part of the effect of pollution on migration runs through wage changes. In such situations, wages and house prices are bad controls, and their inclusion biases the estimated effect of pollution on relocations. Therefore, we need to interpret the estimates with these additional controls with caution.

Table A1 in the Appendix provides the corresponding results. It presents the second-stage estimates with the logarithm of the average wage and a house price index in the origin and destination commuting zones as control variables.[21] Controlling for wages and house prices has little effect on our estimates regarding between-state migration. The coefficients for destination pollution increase in magnitude, which is consistent with a shutdown of the feedback loop. However, the magnitude of the coefficients for origin pollution decreases, which is more in line with a bad control explanation. Including wages and house prices as controls has more pronounced effects on within-state migration, but all estimates remain statistically insignificant. The coefficients for origin wage and destination house prices mostly have an unexpectedly positive sign. However, since wages and house prices are endogenous, we refrain from interpreting their coefficients.

Table 3 present our long-difference estimates. Even more than the annual panel estimates, the long-difference estimates allow us to focus on persistent air quality improvements. Further, they account for the potentially gradual diffusion of information about air quality at destination locations. The results for the long-difference estimates are qualitatively and, in most cases, quantitatively very similar to the ones for the annual panel estimates: In the first stages, simulated origin pollution predicts actual origin pollution and simulated destination pollution predicts actual destination pollution, but not *vice versa*. In the reduced form and second stage, we find positive effects of simulated and actual origin pollution on relocations and negative effects of simulated and actual destination pollution, whereby between-state relocations drive these effects. In contrast to the annual panel estimates, we find slightly larger "pull" than "push" effects with the long-difference estimates. In the full sample, we find "push" and "pull" effects of 0.0034 and −0.0065 per 1,000 households and of 0.0049 and −0.0052 per 1,000 individuals. In contrast to our annual panel results, the pull effects are significant.[22]

To illustrate the magnitude of our estimates, we consider migration responses to an isolated

[21] The data on average wages come from the Bureau of Labor Statistics (https://www.bls.gov/cew/downloadable-data-files.htm, accessed November 24, 2025), the data on house prices from the Federal Housing Finance Agency House Price Index (https://www.fhfa.gov/data/hpi, accessed November 24, 2025). We use the median house price index across counties in a commuting zone, similar to Olney and Thompson (2026).

[22] Table A2 and Table A3 present long-difference estimates with two-year and five-year averages. Focusing on all and between-state relocations, it is mainly the estimates for origin pollution that change in Table A2 compared to Table 3: They become smaller and less statistically significant. In contrast, in Table A3 it is mostly the estimates for destination pollution that become smaller compared to the ones in Table 3 and statistically insignificant.

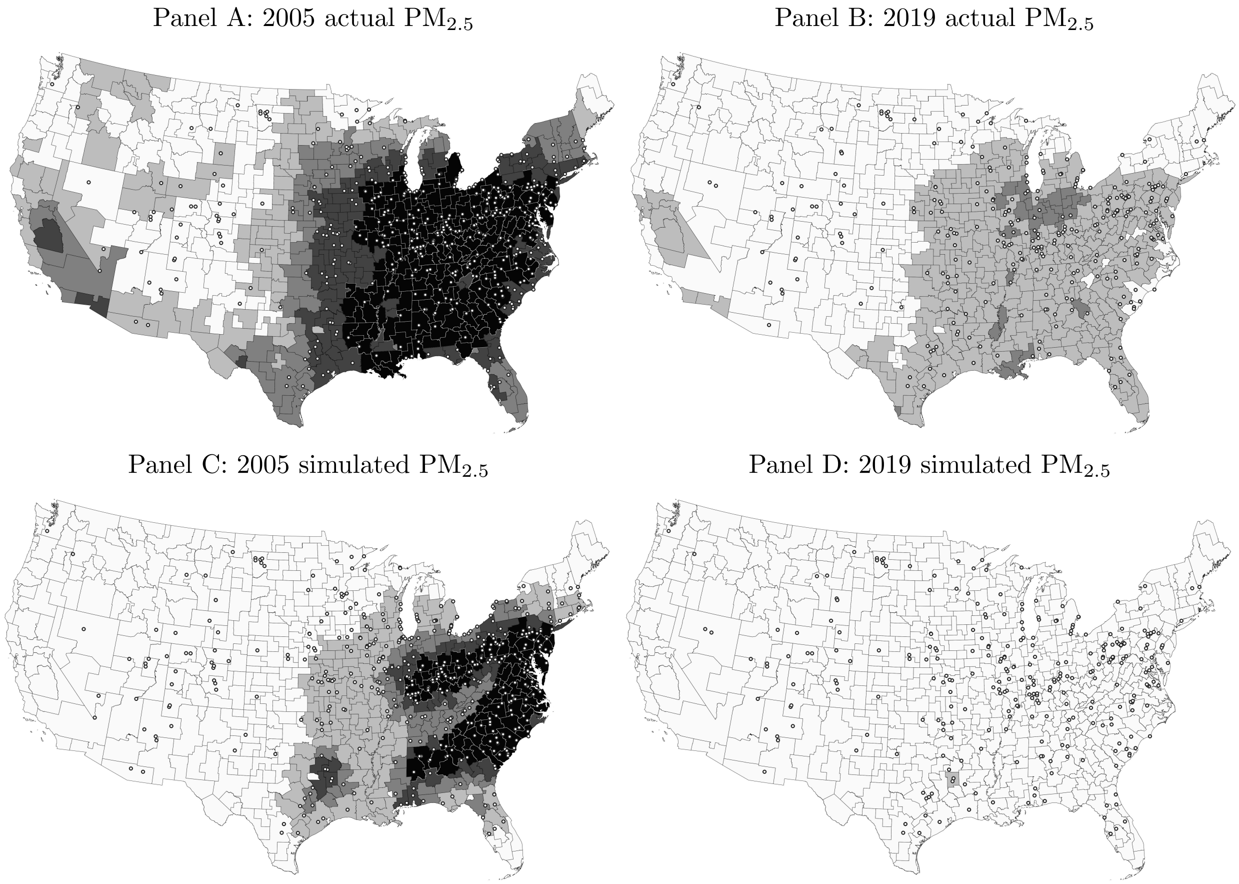
air quality improvement in an average-sized commuting zone of 187,335 households and 408,074 individuals. In our thought experiment, we assume that the $PM_{2.5}$ concentration decreases by 3.85 µg/m^3 – the average reduction observed between 2005 and 2019 – in this representative commuting zone, but remains constant everywhere else. Our coefficient for origin pollution in the household migration regressions of 0.0034 per 1,000 households implies that for a particular directed dyad, the share of migrating households changes by 0.0000034×-3.85 or -0.00001309 due to the hypothetical air quality improvement. In absolute numbers, this amounts to $0.00001309 \times$ 187,335 or 2.45 avoided household relocations. Across all 559 potential destinations, this sums up to an additional 1,370 households that remain in the commuting zone. Similarly, the coefficient for destination pollution of -0.0065 per 1,000 households suggests that the hypothetical air quality improvement increases the share of in-migrating households for a particular directed dyad by -0.0000065×-3.85 or 0.000025025. Since the size of the average origin commuting zone simply corresponds to the size of the average commuting zone, the absolute number of additional households migrating to the representative commuting zone is 4.69 ($= 0.000025025 \times$ 187,335) per directed dyad or 2,622 ($= 4.69 \times 559$) overall. Adding the avoided out-migrations and the additional in-migrations suggests that the hypothetical air quality improvements increase the size of the representative commuting zone by 3,992 (=1,370+2,622) households per year or by 2.13 percent of its initial population. Similar stepwise calculations for individuals yield a population increase for the representative commuting zone of 8,871 individuals per year or 2.17 percent of the initial population.

This thought experiment highlights the importance of air quality for relocation decisions. However, the effects of the actual changes in air pollution on migration patterns will be much smaller. The reason is that air quality improvements are typically not isolated but rather similarly affect origin and destination commuting zones. Thus, the effects of simultaneous improvements at origin and destination locations partly cancel each other out.

5 Conclusion

Two factors complicate an intuitive understanding of the importance of pollution in shaping internal migration patterns. First, the adverse effects of pollution often go hand in hand with pollution-generating activities that attract residents. Second, origin and destination locations typically experience similar changes in air pollution, thereby offsetting their respective effects on migration decisions. However, accounting for the endogeneity of pollution and correlated pollution shocks, we find clear results: Air quality plays an important role in households' and individuals' relocation decisions. The findings are consistent with people seeking to avoid exposure to high pollution levels by changing residence. This insight adds an important component for decision makers pondering the wider social and demographic consequences of decisions affecting regional air quality. More generally, our results contribute to an ever-wider recognition of the pervasive effects of air pollution (Aguilar-Gomez et al. 2022).

Figure 1: 2005 and 2019 actual and simulated $PM_{2.5}$ pollution



Notes: Panels A and B show commuting zone averages of actual $PM_{2.5}$ concentration in µg/m^3 intervals of □ [0, 6), ■ [6, 8), ■ [8, 10), ■ [10, 12), and ■ [12, ∞). Panels C and D depict commuting zone averages of simulated $PM_{2.5}$ concentration in µg/m^3 intervals of □ [0, 2), ■ [2, 4), ■ [4, 6), ■ [6, 8), and ■ [8, ∞). ○ are coal-fired power plants operating in the respective year.

Table 1: Effect of air pollution on relocations, estimates based on annual panel data

	All relocations		Within-state relocations		Between-state relocations	
	Households	Individuals	Households	Individuals	Households	Individuals
Panel A: OLS estimates	*Dependent variable: relocations*					
Origin actual $PM_{2.5}$	-0.0009	-0.0007	-0.0219	-0.0178	-0.0003	-0.0002
	(0.0009)	(0.0008)	(0.0261)	(0.0236)	(0.0003)	(0.0003)
	[0.0008]	[0.0008]	[0.0245]	[0.0220]	[0.0003]	[0.0003]
Destination actual $PM_{2.5}$	0.0008	0.0002	0.0054	-0.0064	0.0006	0.0004
	(0.0005)	(0.0004)	(0.0093)	(0.0065)	(0.0004)	(0.0003)
	[0.0005]	[0.0004]	[0.0094]	[0.0068]	[0.0004]	[0.0003]
Panel B: IV estimates						
First stages	*Dependent variable: origin actual $PM_{2.5}$*					
Origin simulated $PM_{2.5}$	0.1545	0.1545	0.1666	0.1666	0.1541	0.1541
	(0.0291)***	(0.0291)***	(0.0392)***	(0.0392)***	(0.0290)***	(0.0290)***
	[0.0369]***	[0.0369]***	[0.0529]***	[0.0529]***	[0.0367]***	[0.0367]***
Destination simulated $PM_{2.5}$	0.0000	0.0000	0.0008	0.0008	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0014)	(0.0014)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0018]	[0.0019]	[0.0000]	[0.0000]
	Dependent variable: destination actual $PM_{2.5}$					
Origin simulated $PM_{2.5}$	-0.0000	-0.0000	0.0008	0.0008	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0014)	(0.0014)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0020]	[0.0021]	[0.0000]	[0.0000]
Destination simulated $PM_{2.5}$	0.1545	0.1545	0.1666	0.1666	0.1541	0.1541
	(0.0291)***	(0.0291)***	(0.0392)***	(0.0392)***	(0.0290)***	(0.0290)***
	[0.0369]***	[0.0369]***	[0.0529]***	[0.0529]***	[0.0367]***	[0.0367]***
Reduced form	*Dependent variable: relocations*					
Origin simulated $PM_{2.5}$	0.0007	0.0009	0.0010	-0.0003	0.0007	0.0009
	(0.0003)*	(0.0003)***	(0.0084)	(0.0075)	(0.0002)***	(0.0002)***
	[0.0004]*	[0.0003]***	[0.0084]	[0.0075]	[0.0002]***	[0.0002]***
Destination simulated $PM_{2.5}$	-0.0005	-0.0004	0.0010	0.0011	-0.0005	-0.0004
	(0.0005)	(0.0004)	(0.0070)	(0.0059)	(0.0005)	(0.0004)
	[0.0005]	[0.0004]	[0.0067]	[0.0056]	[0.0005]	[0.0004]
Second stage	*Dependent variable: relocations*					
Origin actual $PM_{2.5}$	0.0044	0.0058	0.0058	-0.0016	0.0043	0.0060
	(0.0024)*	(0.0024)**	(0.0504)	(0.0448)	(0.0017)**	(0.0018)***
	[0.0026]*	[0.0025]**	[0.0506]	[0.0449]	[0.0017]**	[0.0019]***
Destination actual $PM_{2.5}$	-0.0032	-0.0026	0.0061	0.0067	-0.0034	-0.0029
	(0.0034)	(0.0028)	(0.0421)	(0.0350)	(0.0031)	(0.0027)
	[0.0035]	[0.0029]	[0.0398]	[0.0330]	[0.0033]	[0.0028]
Controls						
Directed dyad effects	Yes	Yes	Yes	Yes	Yes	Yes
Origin state-by-year effects	Yes	Yes	Yes	Yes	Yes	Yes
Destination state-by-year effects	Yes	Yes	Yes	Yes	Yes	Yes
Weather covariates	Yes	Yes	Yes	Yes	Yes	Yes
Estimation statistics						
Minimum (at the 0.05 level) rejectable 2SLS bias with two-way clustering at						
- commuting zone	14%	14%	36%	36%	14%	14%
- economic area	29%	29%	134%	134%	29%	29%
Commuting zones	560	560	557	557	560	560
Economic areas	175	175	175	175	175	175
Observations	4,695,600	4,695,600	129,690	129,690	4,565,910	4,565,910

Notes: "Relocations" are the number of households or individuals moving from origin to destination per 1,000 households or individuals in the origin. $PM_{2.5}$ is measured in $\mu g/m^3$. Standard errors allowing for two-way clustering at the origin- and destination-commuting zone levels in parentheses and at the origin- and destination-economic area levels in brackets. Three states are matched with only one commuting zone. These commuting zones and states have no observations for within-state relocations. The test of overall instrument strength is from Lewis and Mertens (2025). Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2: Effect of air pollution on environmental concern

	Dependent variable:			
Panel A: OLS estimates	*Concern 1-4*		*Concern: a great deal*	
Actual $PM_{2.5}$	0.0990	0.0834	0.0456	0.0372
	(0.0264)***	(0.0252)***	(0.0137)***	(0.0134)***
	[0.0255]***	[0.0227]***	[0.0132]***	[0.0125]***
Panel B: IV estimates				
First stages	*Dependent variable: actual $PM_{2.5}$*			
Simulated $PM_{2.5}$	0.1403	0.1405	0.1403	0.1405
	(0.0411)***	(0.0410)***	(0.0411)***	(0.0410)***
	[0.0423]***	[0.0422]***	[0.0423]***	[0.0422]***
	Dependent variable:			
Reduced form	*Concern 1-4*		*Concern: a great deal*	
Simulated $PM_{2.5}$	0.0296	0.0427	0.0197	0.0264
	(0.0257)	(0.0285)	(0.0126)	(0.0137)*
	[0.0263]	[0.0296]	[0.0132]	[0.0146]*
	Dependent variable:			
Second stage	*Concern 1-4*		*Concern: a great deal*	
Actual $PM_{2.5}$	0.2112	0.3038	0.1406	0.1882
	(0.2092)	(0.2502)	(0.1107)	(0.1315)
	[0.2133]	[0.2582]	[0.1140]	[0.1371]
Controls				
Commuting zone effects	Yes	Yes	Yes	Yes
State-by-year effects	Yes	Yes	Yes	Yes
Weather covariates	Yes	Yes	Yes	Yes
Demograpics and income	No	Yes	No	Yes
Estimation statistics				
First-stage F statistic				
- clustering at commuting zone	11.68	11.77	11.68	11.77
- clustering at economic area	11.02	11.10	11.02	11.10
Commuting zones	465	465	465	465
Economic areas	173	173	173	173
Observations	12,423	12,423	12,423	12,423

"Concern 1-4" is the original four-point-scale variable on environmental concern, "concern: a great deal" is an indicator for responses in the highest concern category. $PM_{2.5}$ is measured in $\mu g/m^3$. Demographics include age in years and indicators for female, Hispanic, non-White, college graduate, Republican, and Democrat. Income includes indicators for annual income between USD 30,000 and 75,000 and for annual income above USD 75,000. Standard errors allowing for clustering at the commuting zone levels in parentheses and at the economic area levels in brackets. The test of instrument strength is from Kleibergen and Paap (2006). Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 3: Effect of air pollution on relocations, estimates based on long differences

	All relocations		Within-state relocations		Between-state relocations	
	Households	Individuals	Households	Individuals	Households	Individuals
Panel A: OLS estimates	*Dependent variable: relocations*					
Origin actual $PM_{2.5}$	-0.0020	-0.0012	-0.0624	-0.0481	-0.0000	0.0003
	(0.0020)	(0.0018)	(0.0458)	(0.0406)	(0.0010)	(0.0010)
	[0.0019]	[0.0017]	[0.0438]	[0.0393]	[0.0011]	[0.0011]
Destination actual $PM_{2.5}$	0.0025	0.0015	0.0231	-0.0001	0.0016	0.0015
	(0.0015)*	(0.0012)	(0.0272)	(0.0193)	(0.0010)	(0.0009)
	[0.0014]*	[0.0011]	[0.0294]	[0.0210]	[0.0010]	[0.0009]*
Panel B: IV estimates						
First stages	*Dependent variable: origin actual $PM_{2.5}$*					
Origin simul. $PM_{2.5}$	0.1603	0.1603	0.1779	0.1779	0.1595	0.1595
	(0.0289)***	(0.0289)***	(0.0368)***	(0.0368)***	(0.0287)***	(0.0287)***
	[0.0344]***	[0.0344]***	[0.0463]***	[0.0463]***	[0.0343]***	[0.0343]***
Destination simul. $PM_{2.5}$	-0.0000	-0.0000	0.0007	0.0007	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0015)	(0.0015)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0019]	[0.0018]	[0.0000]	[0.0000]
	Dependent variable: destination actual $PM_{2.5}$					
Origin simul. $PM_{2.5}$	0.0000	0.0000	0.0007	0.0007	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0015)	(0.0015)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0019]	[0.0018]	[0.0000]	[0.0000]
Destination simul. $PM_{2.5}$	0.1603	0.1603	0.1779	0.1779	0.1595	0.1595
	(0.0289)***	(0.0289)***	(0.0368)***	(0.0368)***	(0.0287)***	(0.0287)***
	[0.0344]***	[0.0344]***	[0.0463]***	[0.0463]***	[0.0343]***	[0.0343]***
Reduced form	*Dependent variable: relocations*					
Origin simul. $PM_{2.5}$	0.0005	0.0008	-0.0027	-0.0032	0.0006	0.0009
	(0.0005)	(0.0005)*	(0.0089)	(0.0077)	(0.0003)*	(0.0003)***
	[0.0005]	[0.0005]	[0.0096]	[0.0083]	[0.0004]*	[0.0004]**
Destination simul. $PM_{2.5}$	-0.0010	-0.0008	-0.0063	-0.0037	-0.0008	-0.0007
	(0.0006)*	(0.0005)*	(0.0087)	(0.0064)	(0.0005)*	(0.0004)*
	[0.0006]*	[0.0005]*	[0.0090]	[0.0066]	[0.0005]*	[0.0004]*
Second stage	*Dependent variable: relocations*					
Origin actual $PM_{2.5}$	0.0034	0.0049	-0.0152	-0.0178	0.0038	0.0055
	(0.0030)	(0.0028)*	(0.0494)	(0.0428)	(0.0021)*	(0.0021)***
	[0.0033]	[0.0032]	[0.0532]	[0.0464]	[0.0021]*	[0.0022]**
Destination actual $PM_{2.5}$	-0.0065	-0.0052	-0.0351	-0.0208	-0.0053	-0.0044
	(0.0038)*	(0.0030)*	(0.0484)	(0.0357)	(0.0030)*	(0.0026)*
	[0.0037]*	[0.0030]*	[0.0485]	[0.0356]	[0.0031]*	[0.0026]*
Controls						
Origin state effects	Yes	Yes	Yes	Yes	Yes	Yes
Destination state effects	Yes	Yes	Yes	Yes	Yes	Yes
Weather covariates	Yes	Yes	Yes	Yes	Yes	Yes
Estimation statistics						
Minimum (at the 0.05 level) rejectable 2SLS bias with two-way clustering at						
- commuting zone	13%	13%	23%	23%	12%	12%
- economic area	21%	21%	62%	62%	21%	21%
Commuting zones	560	560	557	557	560	560
Economic areas	175	175	175	175	175	175
Observations	313,040	313,040	8,646	8,646	304,394	304,394

Notes: "Relocations" are the number of households or individuals moving from origin to destination per 1,000 households or individuals in the origin. $PM_{2.5}$ is measured in $\mu g/m^3$. Estimated in long differences between the years 2005 and 2019. Standard errors allowing for two-way clustering at the origin- and destination-commuting zone levels in parentheses and at the origin- and destination-economic area levels in brackets. Three states are matched with only one commuting zone. These commuting zones and states have no observations for within-state relocations. The test of overall instrument strength is from Lewis and Mertens (2025). Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

References

Aguilar-Gomez, Sandra, Holt Dwyer, Joshua Graff Zivin, and Matthew Neidell (2022). This is air: The "non-health" effects of air pollution. *Annual Review of Resource Economics 14*(1), 403–425.

Banzhaf, H Spencer and Randall P Walsh (2008). Do people vote with their feet? An empirical test of Tiebout's mechanism. *American Economic Review 98*(3), 843–863.

Bartik, Alexander W, Janet Currie, Michael Greenstone, and Christopher R Knittel (2019). The local economic and welfare consequences of hydraulic fracturing. *American Economic Journal: Applied Economics 11*(4), 105–155.

Bayer, Patrick, Nathaniel Keohane, and Christopher Timmins (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management 58*(1), 1–14.

Baylis, Patrick, Prashant Bharadwaj, Jamie T Mullins, and Nick Obradovich (2025). Climate and migration in the United States. *Journal of Public Economics 249*, 105446.

Borgschulte, Mark, David Molitor, and Eric Yongchen Zou (2024). Air pollution and the labor market: Evidence from wildfire smoke. *Review of Economics and Statistics 106*(6), 1558–1575.

Borusyak, Kirill, Rafael Dix-Carneiro, and Brian Kovak (2022). Understanding migration responses to local shocks. *Available at SSRN 4086847*.

Burney, Jennifer A (2020). The downstream air pollution impacts of the transition from coal to natural gas in the United States. *Nature Sustainability 3*(2), 152–160.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller (2011). Robust inference with multiway clustering. *Journal of Business & Economic Statistics 29*(2), 238–249.

Chang, Tom Y, Joshua Graff Zivin, Tal Gross, and Matthew Neidell (2019). The effect of pollution on worker productivity: Evidence from call center workers in China. *American Economic Journal: Applied Economics 11*(1), 151–172.

Chay, Kenneth Y and Michael Greenstone (2003). The impact of air pollution on infant mortality: Evidence from geographic variation in pollution shocks induced by a recession. *Quarterly Journal of Economics 118*(3), 1121–1167.

Chay, Kenneth Y and Michael Greenstone (2005). Does air quality matter? Evidence from the housing market. *Journal of Political Economy 113*(2), 376–424.

Chen, Shuai, Paulina Oliva, and Peng Zhang (2022). The effect of air pollution on migration: Evidence from China. *Journal of Development Economics 156*, 102833.

Coglianese, John, Todd D Gerarden, and James H Stock (2020). The effects of fuel prices, environmental regulations, and other factors on US coal production, 2008-2016. *Energy Journal 41*(1), 55–82.

Currie, Janet, John Voorheis, and Reed Walker (2023). What caused racial disparities in particulate exposure to fall? New evidence from the Clean Air Act and satellite-based measures of air quality. *American Economic Review 113*(1), 71–97.

Davis, Rebecca J, J Scott Holladay, and Charles Sims (2022). Coal-fired power plant retirements in the United States. *Environmental and Energy Policy and the Economy 3*(1), 4–36.

De Silva, Dakshina G, Anita R Schiller, Aurélie Slechten, and Leonard Wolk (2024). Tiebout sorting and toxic releases. *Environmental and Resource Economics 87*(9), 2487–2520.

Fell, Harrison and Daniel T Kaffine (2018). The fall of coal: Joint impacts of fuel prices and renewables on generation and emissions. *American Economic Journal: Economic Policy 10*(2), 90–116.

Fowler, Christopher S (2024). New Commuting Zone delineation for the US based on 2020 data. *Scientific Data 11*(1), 975.

Fraenkel, Rebecca, Josh Graff Zivin, and Sam Krumholz (2024). The coal transition and its implications for health and housing values. *Land Economics 100*(1), 51–65.

Graff Zivin, Joshua and Matthew Neidell (2012). The impact of pollution on worker productivity. *American Economic Review 102*(7), 3652–3673.

Heblich, Stephan, Alex Trew, and Yanos Zylberberg (2021). East-side story: Historical pollution and persistent neighborhood sorting. *Journal of Political Economy 129*(5), 1508–1552.

Johnsen, Reid, Jacob LaRiviere, and Hendrik Wolff (2019). Fracking, coal, and air quality. *Journal of the Association of Environmental and Resource Economists 6*(5), 1001–1037.

Khanna, Gaurav, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song (2025). The productivity consequences of pollution-induced migration in China. *American Economic Journal: Applied Economics 17*(2), 184–224.

Kleibergen, Frank and Richard Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics 133*(1), 97–126.

Knittel, Christopher R, Konstantinos Metaxoglou, and Andre Trindade (2015). Natural gas prices and coal displacement: Evidence from electricity markets. *NBER Working Paper 21627*.

Lewis, Daniel J and Karel Mertens (2025). A robust test for weak instruments for 2SLS with multiple endogenous regressors. *Review of Economic Studies, forthcoming*.

Mikula, Štěpán and Mariola Pytliková (2026). Migratory responses to air pollution reduction: Evidence from large-scale desulfurization programme. *European Economic Review*, 105154.

Olney, William W and Owen Thompson (2026). The determinants of declining internal migration. *Journal of Urban Economics 153*, 103856.

Shen, Siyuan, Chi Li, Aaron Van Donkelaar, Nathan Jacobs, Chenguang Wang, and Randall V Martin (2024). Enhancing global estimation of fine particulate matter concentrations by including geophysical a priori information in deep learning. *ACS ES&T Air 1*(5), 332–345.

Sprung-Keyser, Ben, Nathaniel Hendren, Sonya Porter, et al. (2022). The radius of economic opportunity: Evidence from migration and local labor markets. *Mimeo, Harvard University*.

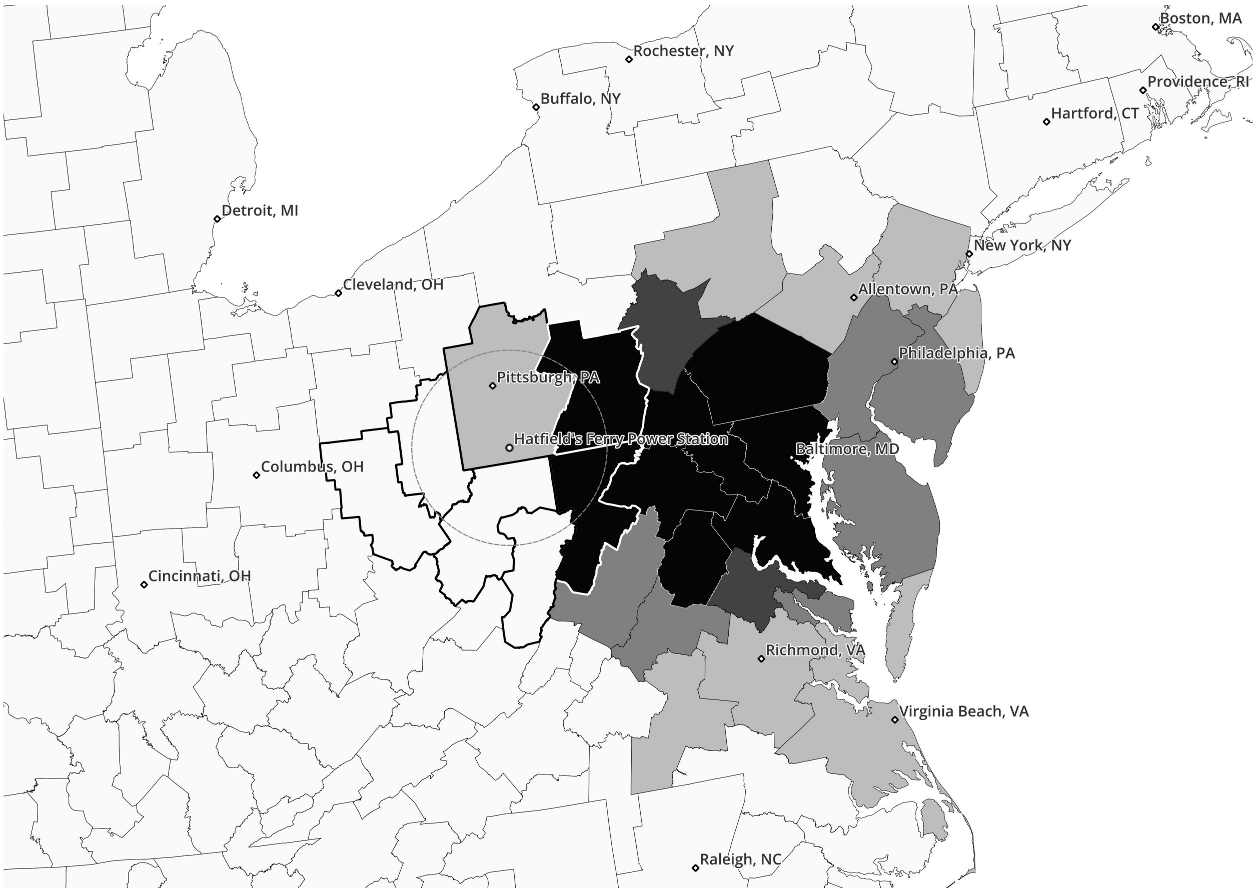
Tessum, Christopher W, Jason D Hill, and Julian D Marshall (2017). InMAP: A model for air pollution interventions. *PloS one 12*(4), e0176131.

Van Donkelaar, Aaron, Melanie S Hammer, Liam Bindle, Michael Brauer, Jeffery R Brook, Michael J Garay, N Christina Hsu, Olga V Kalashnikova, Ralph A Kahn, Colin Lee, et al. (2021). Monthly global estimates of fine particulate matter and their uncertainty. *Environmental Science & Technology 55*(22), 15287–15300.

Wilson, Riley (2022). Moving to economic opportunity: The migration response to the fracking boom. *Journal of Human Resources 57*(3), 918–955.

Appendix. Figures and Tables

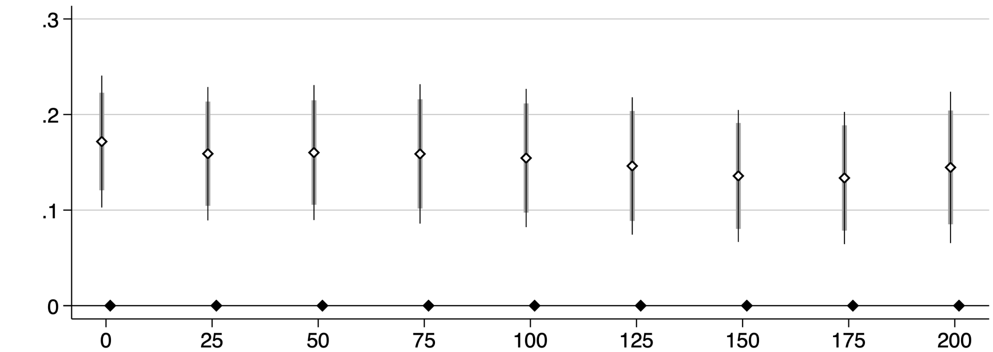
Figure A1: Illustration of pollution simulation and distance threshold



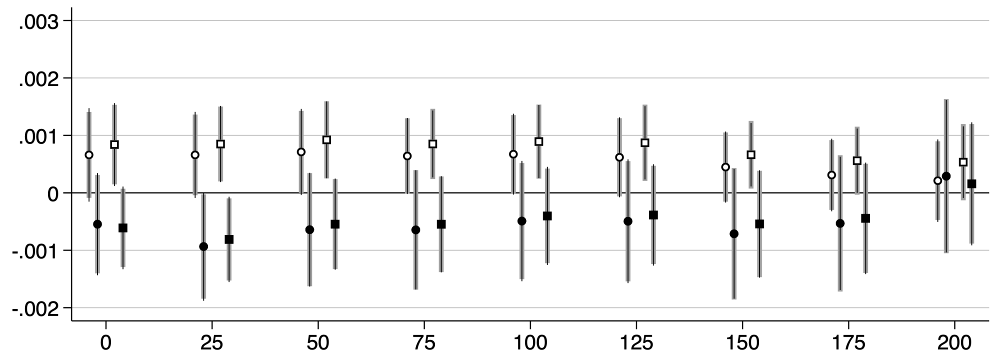
Notes: The figure shows 2005 commuting zone averages of simulated $PM_{2.5}$ concentration originating from Hatfield's Ferry Power Station in µg/m^3 intervals of □ [0, 0.25), ▢ [0.25, 0.5), ▢ [0.5, 0.75), ■ [0.75, 1), and ■ [1, ∞). $\circ$ is Hatfield's Ferry Power Station. $\diamond$ are cities with a 2019 population of at least 100,000. The figure also includes a 100-kilometer-radius circle centered on the plant. Commuting zones with thick border lines are those intersecting with the circle area.

Figure A2: Sensitivity test for all relocations: varying distance threshold

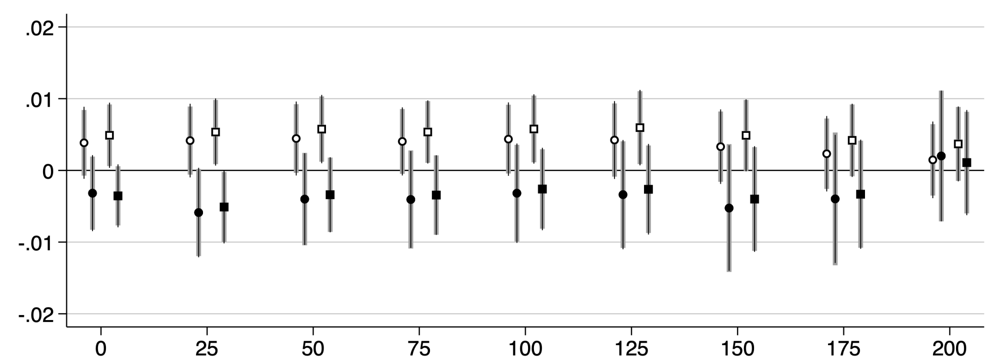
Panel A: first stage (dependent variable: origin actual $PM_{2.5}$)



Panel B: reduced form (dependent variable: relocations)

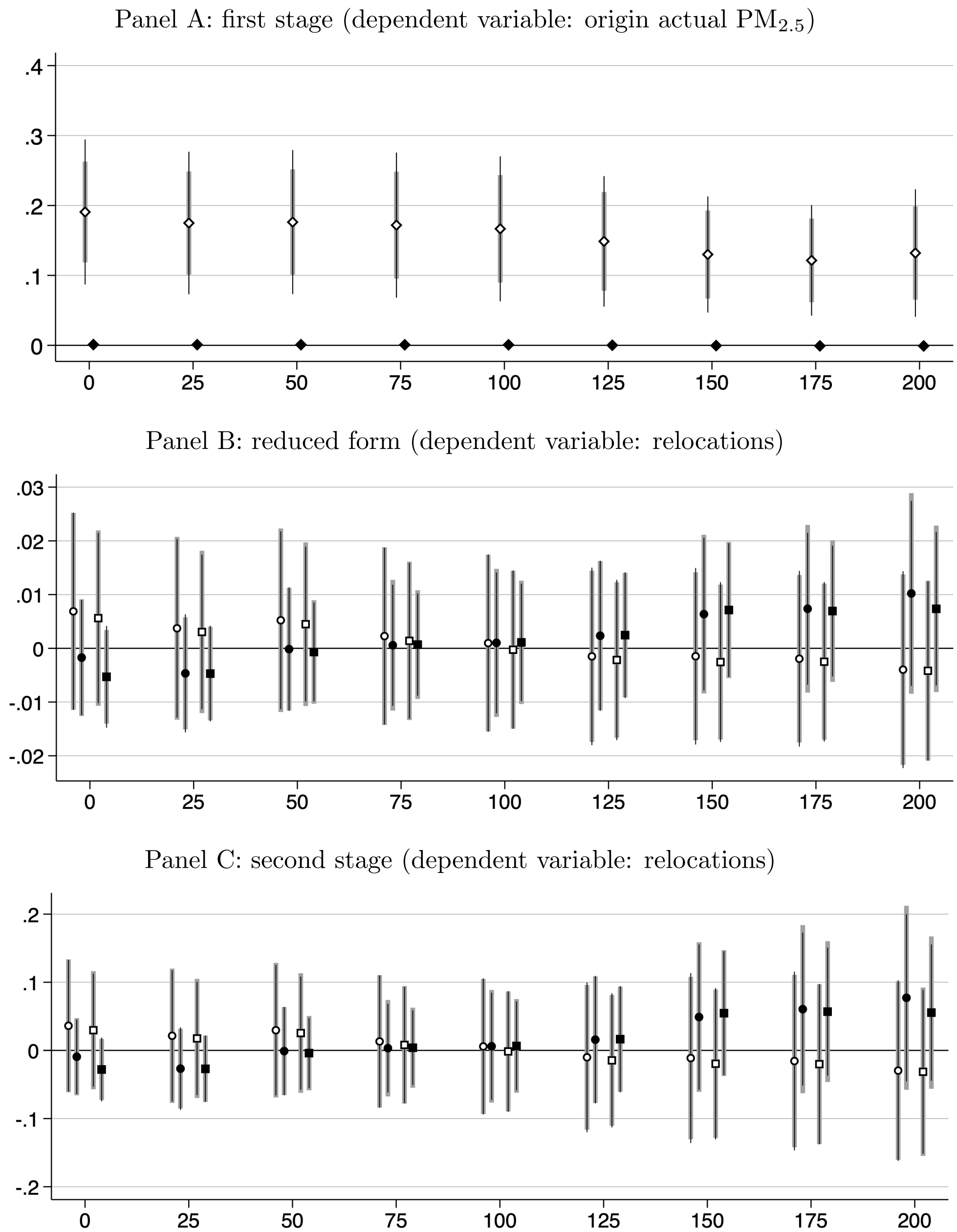


Panel C: second stage (dependent variable: relocations)



Notes: The y-axes show point estimates (symbols) and 95-percent confidence intervals constructed from standard errors allowing for clustering at the origin- and destination-commuting zone levels (thick line) and at the origin- and destination-economic area levels (thin line). Hollow symbols are for origin variables, filled for destination variables. The x-axes depict different distance thresholds. These refer to the minimum distance between a power plant and a commuting zone's boundary for the power plant's simulated pollution to be counted as contributing to the commuting zone's pollution. Panel A depicts the first-stage estimates from regressing actual origin pollution on simulated origin pollution (◇) and simulated destination pollution (◆). We do not show the first-stage estimates for actual destination pollution, as these are symmetric to the ones for actual origin pollution. Panel B shows the reduced-form estimates from regressing household relocations on simulated origin pollution (○) and simulated destination pollution (●) and from regressing individual relocations on simulated origin pollution (□) and simulated destination pollution (■). Panel C presents the corresponding second-stage estimates.

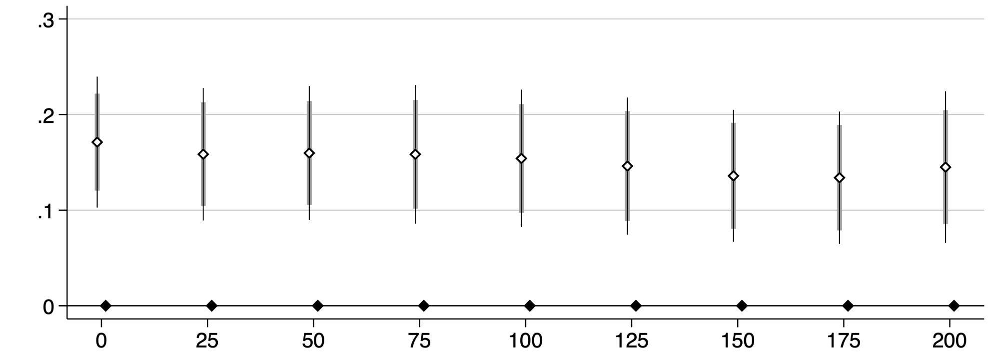
Figure A3: Sensitivity test for within-state relocations: varying distance threshold



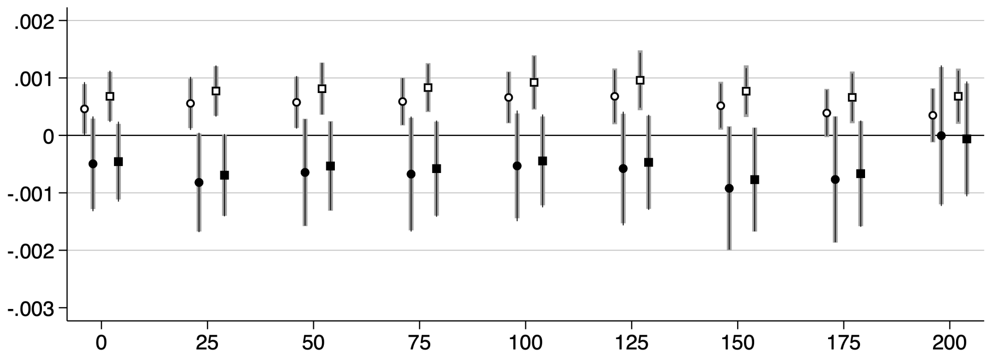
Notes: The y-axes show point estimates (symbols) and 95-percent confidence intervals constructed from standard errors allowing for clustering at the origin- and destination-commuting zone levels (thick line) and at the origin- and destination-economic area levels (thin line). Hollow symbols are for origin variables, filled for destination variables. The x-axes depict different distance thresholds. These refer to the minimum distance between a power plant and a commuting zone's boundary for the power plant's simulated pollution to be counted as contributing to the commuting zone's pollution. Panel A depicts the first-stage estimates from regressing actual origin pollution on simulated origin pollution (◇) and simulated destination pollution (◆). We do not show the first-stage estimates for actual destination pollution, as these are symmetric to the ones for actual origin pollution. Panel B shows the reduced-form estimates from regressing household relocations on simulated origin pollution (○) and simulated destination pollution (●) and from regressing individual relocations on simulated origin pollution (□) and simulated destination pollution (■). Panel C presents the corresponding second-stage estimates.

Figure A4: Sensitivity test for between-state relocations: varying distance threshold

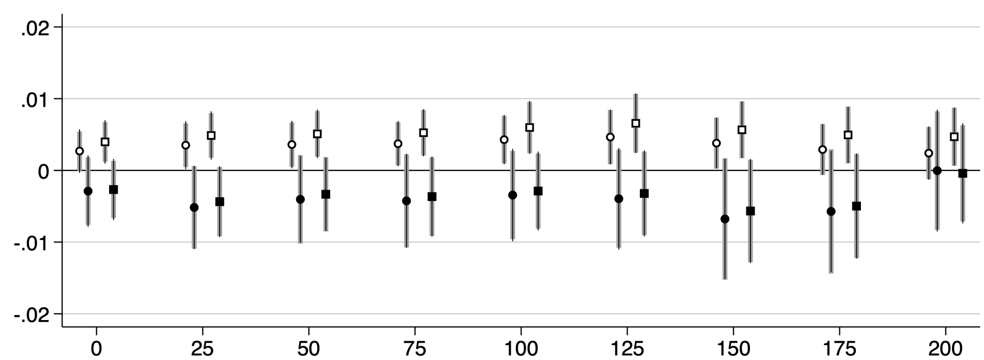
Panel A: first stage (dependent variable: origin actual $PM_{2.5}$)



Panel B: reduced form (dependent variable: relocations)



Panel C: second stage (dependent variable: relocations)



Notes: The y-axes show point estimates (symbols) and 95-percent confidence intervals constructed from constructed from allowing for clustering at the origin- and destination-commuting zone levels (thick line) and at the origin- and destination-economic area levels (thin line). Hollow symbols are for origin variables, filled for destination variables. The x-axes depict different distance thresholds. These refer to the minimum distance between a power plant and a commuting zone's boundary for the power plant's simulated pollution to be counted as contributing to the commuting zone's pollution. Panel A depicts the first-stage estimates from regressing actual origin pollution on simulated origin pollution (◇) and simulated destination pollution (◆). We do not show the first-stage estimates for actual destination pollution, as these are symmetric to the ones for actual origin pollution. Panel B shows the reduced-form estimates from regressing household relocations on simulated origin pollution (○) and simulated destination pollution (●) and from regressing individual relocations on simulated origin pollution (□) and simulated destination pollution (■). Panel C presents the corresponding second-stage estimates.

Table A1: Sensitivity test: controlling for wages and house prices

	All relocations		Within-state relocations		Between-state relocations	
	Households	Individuals	Households	Individuals	Households	Individuals
Second stage	*Dependent variable: relocations*					
Origin actual $PM_{2.5}$	0.0026	0.0040	-0.0363	-0.0446	0.0037	0.0054
	(0.0021)	(0.0020)*	(0.0381)	(0.0341)	(0.0017)**	(0.0018)***
	[0.0021]	[0.0020]**	[0.0385]	[0.0336]	[0.0017]**	[0.0018]***
Origin wage	0.0397	0.0382	0.0803	0.0784	0.0403	0.0385
	(0.0198)**	(0.0186)**	(0.1869)	(0.1715)	(0.0179)**	(0.0170)**
	[0.0242]	[0.0226]*	[0.1977]	[0.1866]	[0.0217]*	[0.0205]*
Origin HPI	0.0115	0.0123	0.1753	0.2217	0.0073	0.0068
	(0.0049)**	(0.0044)***	(0.0819)**	(0.0751)***	(0.0040)*	(0.0037)*
	[0.0049]**	[0.0044]***	[0.0824]**	[0.0760]***	[0.0041]*	[0.0039]*
Destination actual $PM_{2.5}$	-0.0038	-0.0028	-0.0104	0.0019	-0.0038	-0.0031
	(0.0035)	(0.0028)	(0.0463)	(0.0379)	(0.0031)	(0.0027)
	[0.0035]	[0.0029]	[0.0437]	[0.0354]	[0.0033]	[0.0028]
Destination wage	0.0385	0.0302	0.7488	0.5302	0.0167	0.0147
	(0.0102)***	(0.0076)***	(0.2375)***	(0.1661)***	(0.0062)***	(0.0053)***
	[0.0108]***	[0.0082]***	[0.2398]***	[0.1726]***	[0.0063]***	[0.0054]***
Destination HPI	0.0047	-0.0002	0.1391	0.0121	0.0019	0.0003
	(0.0062)	(0.0043)	(0.1585)	(0.1076)	(0.0037)	(0.0028)
	[0.0067]	[0.0046]	[0.1782]	[0.1190]	[0.0036]	[0.0027]
Controls						
Directed dyad effects	Yes	Yes	Yes	Yes	Yes	Yes
Origin state-by-year effects	Yes	Yes	Yes	Yes	Yes	Yes
Destination state-by-year effects	Yes	Yes	Yes	Yes	Yes	Yes
Weather covariates	Yes	Yes	Yes	Yes	Yes	Yes
Estimation statistics						
Minimum (at the 0.05 level) rejectable 2SLS bias with two-way clustering at						
- commuting zone	13%	13%	32%	32%	13%	13%
- economic area	27%	27%	114%	114%	27%	27%
Commuting zones	545	545	542	542	545	545
Economic areas	175	175	175	175	175	175
Observations	4,447,200	4,447,200	118,800	118,800	4,328,400	4,328,400

Notes: "Relocations" are the number of households or individuals moving from origin to destination per 1,000 households or individuals in the origin. $PM_{2.5}$ is measured in $\mu g/m^3$. Wages are the natural logarithm of the average commuting zone wage. House prices are the natural logarithm of the median house price index across counties in a commuting zone. Standard errors allowing for two-way clustering at the origin- and destination-commuting zone levels in parentheses and at the origin- and destination-economic area levels in brackets. Three states are matched with only one commuting zone. These commuting zones and states have no observations for within-state relocations. We have fewer observations than in Table 1 due to missing house price data. The test of overall instrument strength is from Lewis and Mertens (2025). Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A2: Effect of air pollution on relocations, estimates based on long differences using two-year averages

	All relocations		Within-state relocations		Between-state relocations	
	Households	Individuals	Households	Individuals	Households	Individuals
Panel A: OLS estimates			*Dependent variable: relocations*			
Origin actual $PM_{2.5}$	-0.0036	-0.0031	-0.0718	-0.0639	-0.0013	-0.0010
	(0.0026)	(0.0024)	(0.0668)	(0.0600)	(0.0008)*	(0.0007)
	[0.0024]	[0.0021]	[0.0619]	[0.0554]	[0.0008]*	[0.0007]
Destination actual $PM_{2.5}$	0.0036	0.0023	0.0478	0.0185	0.0020	0.0017
	(0.0014)**	(0.0011)**	(0.0289)*	(0.0206)	(0.0009)**	(0.0008)**
	[0.0013]***	[0.0010]**	[0.0320]	[0.0230]	[0.0008]**	[0.0007]**
Panel B: IV estimates						
First stages			*Dependent variable: origin actual $PM_{2.5}$*			
Origin simul. $PM_{2.5}$	0.1457	0.1457	0.1598	0.1598	0.1452	0.1452
	(0.0322)***	(0.0322)***	(0.0427)***	(0.0427)***	(0.0320)***	(0.0320)***
	[0.0411]***	[0.0411]***	[0.0567]***	[0.0567]***	[0.0409]***	[0.0409]***
Destination simul. $PM_{2.5}$	-0.0000	-0.0000	0.0007	0.0007	0.0000	0.0000
	(0.0001)	(0.0001)	(0.0016)	(0.0016)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0021]	[0.0021]	[0.0000]	[0.0000]
			Dependent variable: destination actual $PM_{2.5}$			
Origin simul. $PM_{2.5}$	-0.0000	-0.0000	0.0007	0.0007	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0017)	(0.0016)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0021]	[0.0021]	[0.0000]	[0.0000]
Destination simul. $PM_{2.5}$	0.1457	0.1457	0.1598	0.1598	0.1452	0.1452
	(0.0322)***	(0.0322)***	(0.0427)***	(0.0427)***	(0.0320)***	(0.0320)***
	[0.0411]***	[0.0411]***	[0.0567]***	[0.0567]***	[0.0409]***	[0.0409]***
Reduced form			*Dependent variable: relocations*			
Origin simul. $PM_{2.5}$	0.0004	0.0005	0.0018	0.0001	0.0004	0.0006
	(0.0004)	(0.0004)	(0.0088)	(0.0079)	(0.0002)	(0.0002)**
	[0.0004]	[0.0004]	[0.0094]	[0.0084]	[0.0002]	[0.0002]***
Destination simul. $PM_{2.5}$	-0.0009	-0.0007	0.0008	0.0014	-0.0008	-0.0007
	(0.0006)	(0.0005)	(0.0085)	(0.0066)	(0.0005)*	(0.0004)*
	[0.0006]	[0.0005]	[0.0082]	[0.0065]	[0.0005]*	[0.0004]*
Second stage			*Dependent variable: relocations*			
Origin actual $PM_{2.5}$	0.0026	0.0038	0.0115	0.0004	0.0026	0.0040
	(0.0028)	(0.0026)	(0.0554)	(0.0494)	(0.0018)	(0.0019)**
	[0.0030]	[0.0028]	[0.0590]	[0.0523]	[0.0018]	[0.0020]**
Destination actual $PM_{2.5}$	-0.0059	-0.0047	0.0048	0.0090	-0.0058	-0.0048
	(0.0040)	(0.0032)	(0.0529)	(0.0411)	(0.0034)*	(0.0029)*
	[0.0040]	[0.0033]	[0.0509]	[0.0403]	[0.0034]*	[0.0029]
Controls						
Origin state effects	Yes	Yes	Yes	Yes	Yes	Yes
Destination state effects	Yes	Yes	Yes	Yes	Yes	Yes
Weather covariates	Yes	Yes	Yes	Yes	Yes	Yes
Estimation statistics						
Minimum (at the 0.05 level) rejectable 2SLS bias with two-way clustering at						
- commuting zone	23%	23%	58%	58%	22%	22%
- economic area	51%	51%	236%	236%	51%	51%
Commuting zones	560	560	557	557	560	560
Economic areas	175	175	175	175	175	175
Observations	313,040	313,040	8,646	8,646	304,394	304,394

Notes: "Relocations" are the number of households or individuals moving from origin to destination per 1,000 households or individuals in the origin. $PM_{2.5}$ is measured in $\mu g/m^3$. Estimated in long differences between the 2005-2006 and 2018-2019 averages. Standard errors allowing for two-way clustering at the origin- and destination-commuting zone levels in parentheses and at the origin- and destination-economic area levels in brackets. Three states are matched with only one commuting zone. These commuting zones and states have no observations for within-state relocations. The test of overall instrument strength is from Lewis and Mertens (2025). Significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table A3: Effect of air pollution on relocations, estimates based on five-year long differences using five-year averages

	All relocations		Within-state relocations		Between-state relocations	
	Households	Individuals	Households	Individuals	Households	Individuals
Panel A: OLS estimates	*Dependent variable: relocations*					
Origin actual $PM_{2.5}$	-0.0039	-0.0033	-0.0884	-0.0831	-0.0010	-0.0006
	(0.0031)	(0.0029)	(0.0806)	(0.0767)	(0.0007)	(0.0007)
	[0.0028]	[0.0027]	[0.0736]	[0.0707]	[0.0007]	[0.0007]
Destination actual $PM_{2.5}$	0.0018	0.0002	0.0096	-0.0231	0.0015	0.0010
	(0.0011)	(0.0009)	(0.0223)	(0.0163)	(0.0008)*	(0.0007)
	[0.0012]	[0.0009]	[0.0231]	[0.0174]	[0.0008]*	[0.0007]
Panel B: IV estimates						
First stages	*Dependent variable: origin actual $PM_{2.5}$*					
Origin simul. $PM_{2.5}$	0.1542	0.1542	0.1744	0.1744	0.1535	0.1535
	(0.0315)***	(0.0315)***	(0.0435)***	(0.0435)***	(0.0313)***	(0.0313)***
	[0.0404]***	[0.0404]***	[0.0595]***	[0.0595]***	[0.0402]***	[0.0402]***
Destination simul. $PM_{2.5}$	-0.0000	-0.0000	0.0012	0.0012	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0015)	(0.0016)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0022]	[0.0022]	[0.0000]	[0.0000]
	Dependent variable: destination actual $PM_{2.5}$					
Origin simul. $PM_{2.5}$	0.0000	0.0000	0.0012	0.0012	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0016)	(0.0016)	(0.0000)	(0.0000)
	[0.0001]	[0.0001]	[0.0022]	[0.0022]	[0.0000]	[0.0000]
Destination simul. $PM_{2.5}$	0.1542	0.1542	0.1744	0.1744	0.1535	0.1535
	(0.0315)***	(0.0315)***	(0.0435)***	(0.0435)***	(0.0313)***	(0.0313)***
	[0.0404]***	[0.0404]***	[0.0595]***	[0.0595]***	[0.0402]***	[0.0402]***
Reduced form	*Dependent variable: relocations*					
Origin simul. $PM_{2.5}$	0.0005	0.0008	-0.0051	-0.0057	0.0007	0.0009
	(0.0003)	(0.0003)**	(0.0089)	(0.0083)	(0.0002)***	(0.0002)***
	[0.0003]	[0.0003]**	[0.0092]	[0.0085]	[0.0002]***	[0.0002]***
Destination simul. $PM_{2.5}$	-0.0005	-0.0003	-0.0022	-0.0006	-0.0004	-0.0003
	(0.0005)	(0.0005)	(0.0074)	(0.0065)	(0.0005)	(0.0004)
	[0.0005]	[0.0005]	[0.0070]	[0.0062]	[0.0005]	[0.0004]
Second stage	*Dependent variable: relocations*					
Origin actual $PM_{2.5}$	0.0035	0.0050	-0.0289	-0.0325	0.0044	0.0060
	(0.0024)	(0.0024)**	(0.0493)	(0.0460)	(0.0017)**	(0.0019)***
	[0.0025]	[0.0024]**	[0.0501]	[0.0460]	[0.0019]**	[0.0020]***
Destination actual $PM_{2.5}$	-0.0032	-0.0022	-0.0126	-0.0034	-0.0028	-0.0021
	(0.0036)	(0.0031)	(0.0427)	(0.0370)	(0.0034)	(0.0029)
	[0.0036]	[0.0031]	[0.0402]	[0.0353]	[0.0034]	[0.0029]
Controls						
Origin state effects	Yes	Yes	Yes	Yes	Yes	Yes
Destination state effects	Yes	Yes	Yes	Yes	Yes	Yes
Weather covariates	Yes	Yes	Yes	Yes	Yes	Yes
Estimation statistics						
Minimum (at the 0.05 level) rejectable 2SLS bias with two-way clustering at						
- commuting zone	18%	18%	45%	45%	18%	18%
- economic area	39%	39%	191%	191%	39%	39%
Commuting zones	560	560	557	557	560	560
Economic areas	175	175	175	175	175	175
Observations	313,040	313,040	8,646	8,646	304,394	304,394

Notes: "Relocations" are the number of households or individuals moving from origin to destination per 1,000 households or individuals in the origin. $PM_{2.5}$ is measured in $\mu g/m^3$. Estimated in long differences between the 2005-2009 and 2015-2019 averages. Standard errors allowing for two-way clustering at the origin- and destination-commuting zone levels in parentheses and at the origin- and destination-economic area levels in brackets. Three states are matched with only one commuting zone. These commuting zones and states have no observations for within-state relocations. The test of overall instrument strength is from Lewis and Mertens (2025). Significance: *** p < 0.01, ** p < 0.05, * p < 0.1.